Environmental Monitoring and Assessment (2006) **121:** 571–596
DOI: 10.1007/s10661-005-9156-7

PATTERNS OF SPATIAL AUTOCORRELATION IN STREAM WATER CHEMISTRY

ERIN E. PETERSON[1,*], ANDREW A. MERTON[2], DAVID M. THEOBALD[3] and N. SCOTT URQUHART[2]

[1]*CSIRO Mathematical & Information Sciences, Queensland Bioscience Precinct, 306 Carmody Road, St. Lucia, QLD, 4067, Australia;* [2]*Department of Statistics, Colorado State University, Fort Collins, CO, USA;* [3]*Natural Resource Ecology Lab and Natural Resources Recreation and Tourism Department Colorado State University, Fort Collins, CO, U.S.A.*

*(*author for correspondence, e-mail: Erin.Peterson@csiro.au)*

(Received 15 July 2005; accepted 2 December 2005)

Abstract. Geostatistical models are typically based on symmetric straight-line distance, which fails to represent the spatial configuration, connectivity, directionality, and relative position of sites in a stream network. Freshwater ecologists have explored spatial patterns in stream networks using hydrologic distance measures and new geostatistical methodologies have recently been developed that enable directional hydrologic distance measures to be considered. The purpose of this study was to quantify patterns of spatial correlation in stream water chemistry using three distance measures: straight-line distance, symmetric hydrologic distance, and weighted asymmetric hydrologic distance. We used a dataset collected in Maryland, USA to develop both general linear models and geostatistical models (based on the three distance measures) for acid neutralizing capacity, conductivity, pH, nitrate, sulfate, temperature, dissolved oxygen, and dissolved organic carbon. The spatial AICC methodology allowed us to fit the autocorrelation and covariate parameters simultaneously and to select the model with the most support in the data. We used the universal kriging algorithm to generate geostatistical model predictions. We found that spatial correlation exists in stream chemistry data at a relatively coarse scale and that geostatistical models consistently improved the accuracy of model predictions. More than one distance measure performed well for most chemical response variables, but straight-line distance appears to be the most suitable distance measure for regional geostatistical modeling. It may be necessary to develop new survey designs that more fully capture spatial correlation at a variety of scales to improve the use of weighted asymmetric hydrologic distance measures in regional geostatistical models.

Keywords: geostatistics, hydrologic distance, scale, spatial autocorrelation, stream networks, water chemistry, weighted asymmetric hydrologic distance

1. Introduction

Stream water chemistry is spatially and temporally heterogeneous at multiple scales (Pringle, 1991; Chambers *et al.*, 1992; Dawson *et al.*, 2001) and the conditions observed at survey sites result from the collective influence of multi-scale landscape filters (Frissell *et al.*, 1986; Poff, 1997). A hierarchical constraint exists among filters (Frissell *et al.*, 1986; Davies *et al.*, 2000), but the strength of the linkage varies between filter scales. Processes acting across spatial and temporal filter scales

produce spatial patterns in water chemistry (Poff, 1997). The quantification of these patterns provides information concerning the importance of ecosystem processes and spatial relationships occurring across multiple scales.

Geostatistical models are commonly used to quantify spatial patterns in the terrestrial environment, but have been applied less frequently to aquatic systems such as lakes (Altunkaynak *et al.*, 2003), estuaries (Little *et al.*, 1997; Rathbun, 1998), and streams (Kellum, 2003; Yuan, 2004). Geostatistical models are typically based on symmetric straight-line distance, which may not be an ecologically representative distance measure because it fails to represent the spatial configuration, connectivity, directionality, and relative position of sites in a stream network (Olden *et al.*, 2001; Benda *et al.*, 2004; Ganio *et al.*, 2005). Recently, freshwater ecologists have begun to explore spatial patterns in stream networks using hydrologic distance measures (Dent and Grimm, 1999; Gardner *et al.*, 2003; Legleiter *et al.*, 2003; Torgersen *et al.*, 2004; Ganio *et al.*, 2005). In addition, new geostatistical methodologies have recently been developed that enable directional hydrologic distance measures to be considered (Ver Hoef *et al.*, 2007). This provides freshwater ecologists with a variety of distance measures to choose from, but it is not obvious which measure is most appropriate for water chemistry data.

Our ability to detect patterns of spatial correlation depends on the grain of the survey design, the extent of the study area, and the configuration of survey locations (Levin, 1992; Cooper *et al.*, 1997). It is likely that the distance measure will also affect the spatial patterns that we observe. Different patterns are likely to occur within and between filter scales (Poff, 1997) and freshwater ecologists must choose the survey scale, design, and distance measure that is most appropriate for their research questions. Little information is available concerning the effect of the distance measure on observed patterns of spatial correlation in stream networks (but see Gardner *et al.*, 2003). Therefore, the purpose of this study is to explore and quantify patterns of spatial correlation in chemical response variables using three distance measures: straight-line distance (SLD), symmetric hydrologic distance (SHD), and weighted asymmetric hydrologic distance (WAHD).

1.1. BACKGROUND

Symmetric and asymmetric distances can be used to represent physical and ecological processes in stream ecosystems. These distances are in turn used to characterize the spatial neighborhood for each site. A spatial neighborhood includes sites that are nearby and have a quantifiable influence upon one another, i.e., are correlated with one another. Sites outside of the spatial neighborhood are considered spatially uncorrelated. Symmetric distance is directionless (isotropic) and has equal correlation in all directions (or both on a stream). SLD is symmetric and all locations in a study area can be considered neighbors (Figure 1a). *Hydrologic distance* can be either symmetric or asymmetric and is simply the distance between two locations when movement is restricted to the stream network. SHD is the total upstream

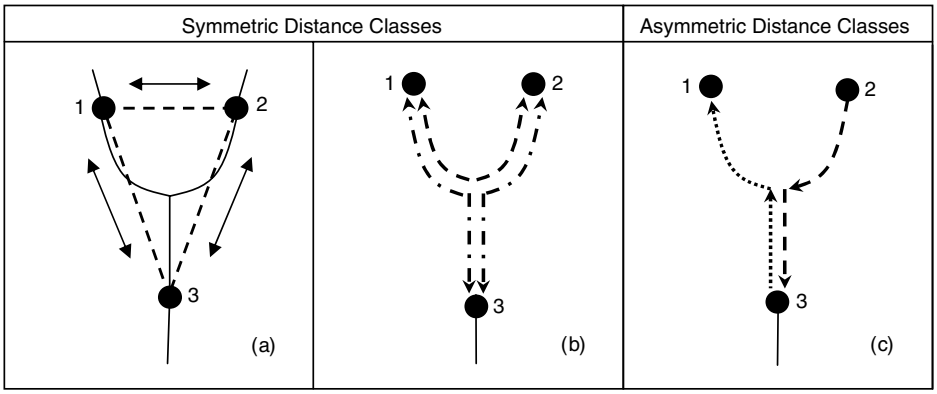

Figure 1. Symmetric and asymmetric distance classes. The stream network is represented by a solid line, while distance measurements are represented with dotted lines. Symmetric hydrologic distance measures include straight-line distance (a) and symmetric hydrologic distance (b). Sites 1, 2, and 3 are all neighbors to one another when these distance measures are used. Asymmetric distance classes include upstream and downstream asymmetric hydrologic distance (c). Sites 1 and 2 are neighbors to site 3, but not to each other.

and downstream hydrologic distance between two sites (Figure 1b). Thus, all sites located within a stream network are neighboring sites (assuming all basins in the study area share a common outlet) because flow direction is disregarded. Asymmetric distances include unidirectional measures that are restricted to either the upstream or downstream flow direction. Water must flow from one location to another to be considered neighbors (Figure 1c). Spatial weights can also be generated using metrics that represent relative network position, such as watershed area, and used to create more ecologically representative WAHD measures (Ver Hoef *et al.*, 2007).

Terrestrial processes may be better represented with SLD because the terrestrial landscape is represented as a two-dimensional surface where any two sites may be connected. For example, simplistic models representing terrestrial transport mechanisms, such as seed dispersal, provide few restrictions to the direction of movement. At times, it may also be appropriate to apply SLD to stream ecosystems. For example, a chemical response variable may be significantly influenced by a continuous landscape variable, such as geology type (Kellum, 2003) or by broad-scale factors, such as acid precipitation (Driscoll *et al.*, 2001). However, SLD may not be as useful when instream processes dominate water chemistry conditions. Although freshwater systems have four dimensions (Ward, 1989), we focus here on the longitudinal or upstream-downstream dimension and represent streams as one-dimensional features. Freshwater riverine systems differ from terrestrial ecosystems because they are linear and typically the movement of material is restricted to the stream network. Some aquatic fauna (e.g. fish) move both up and downstream, but cannot move across the terrestrial landscape (Colyer *et al.*, 2005). Water chemistry is

strongly influenced by longitudinal transport mechanisms. Movement is restricted to the network and occurs primarily in the downstream direction (Closs *et al.*, 2004).

The relative position in the network also affects the condition of a site (Pringle, 2001; Benda *et al.*, 2004) and reflects the influence that it will have on other sites (Cumming, 2002). For example, a site located on a small tributary may have little influence on a downstream site located on the mainstem due to substantial differences in discharge volume (Benda *et al.*, 2004). Clearly, both topographical and topological characteristics of the stream network provide a vast amount of information about chemical conditions at unobserved sites. Therefore, functional distances based on hydrologic connectivity should be considered for geostatistical modeling in stream networks.

Patterns of spatial correlation are visualized using a graphical representation called an empirical semivariogram, which is a plot of the semivariance between sites given their separation distance. The semivariance represents the strength of spatial correlation between two sites (Olea, 1991) and the separation distance is simply the distance traveled from one location to a second location. Semivariograms are generated by dividing the separation distances into groups, or bins, calculating the mean semivariance for each bin (1), and plotting the semivariances for the bins in ascending order (i.e. 100, 200, 300 . . . .). The semivariance is given by

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z_i - Z_{i+h}]^2 , \tag{1}$$

where h is the mean separation distance between sites within a bin, $\gamma(h)$ is the semivariance for the bin, Z_i is the observed sample value at site i, Z_{i+h} is the observed sample value at $i + h$, and $N(h)$ is the total number of sample pairs for the bin.

The semivariogram is a powerful diagnostic tool for identifying whether there is correlation among sites. It is common to fit an autocorrelation function, such as the exponential function, to the semivariogram in order to estimate the three autocorrelation parameters (θ): the nugget, sill, and range (see Gardner *et al.*, 2003; Kellum, 2003; Yuan, 2004). The nugget represents the variation between sites as their separation distance approaches zero. It can result from experimental error or could indicate that a substantial amount of variation occurs at a scale finer than the survey scale. The sill is delineated where the autocorrelation asymptotes and represents variance found among uncorrelated data. The range parameter describes how fast the autocorrelation decays with distance.

Fitting an autocorrelation function to the empirical semivariogram using a least squares method provides robust estimates of the autocorrelation parameters regardless of the error distribution (Carroll and Ruppert, 1982). However, a drawback to this method is that the investigator must select bin size in order to generate a useful semivariogram. The fitted values of the autocorrelation parameters are, therefore,

dependent on the bin size selected. Thus, parameter estimates can vary from investigator to investigator as a function of bin size.

Likelihood-based approaches can also be used to obtain estimates of the autocorrelation parameters (Kitanidis, 1983) and are slightly more robust when the errors are normally distributed (Jobson and Fuller, 1980). However, they are sensitive to misspecification of the distribution and may provide poor parameter estimates when the assumption of normality is violated (Carroll and Ruppert, 1982). The log likelihood function is used to estimate unknown parameter values based on the data and is derived from a specific probability distribution function, such as the exponential autocorrelation function (Hoeting *et al.*, 2006). The data values and the probability distribution function are fixed while the parameter values are allowed to vary. Maximizing the (log) likelihood with respect to the unknown parameters provides maximum likelihood (ML) estimates. ML estimation is an efficient method of parameter estimation for large sample sizes and provides a means for estimating uncertainty in the estimates (Pardo-Iğuzquiza, 1998). In addition, competing models can be readily compared using, for example, the Corrected Akaike's Information Criterion (AICC), which is itself likelihood based (Akaike, 1973).

Few studies have explored patterns of spatial correlation in stream chemistry data (but see Dent and Grimm, 1999; Gardner *et al.*, 2003; Kellum, 2003; Yuan, 2004), yet findings suggest that the range varies with respect to the distance measure, spatial correlation differs between chemical response variables, and that patterns of spatial correlation change over time (Table I). Gardner and others (2003) compared spatial correlation in temperature using three distance measures: SLD, SHD, and SHD weighted by Strahler stream order. They found that hydrologic distance measures led to semivariograms with a larger range than SLD and that weighting hydrologic distance further increased the range (Table I). This was surprising since theoretically SLD would explain patterns of spatial correlation produced by broad-scale ecological processes that are not constrained to one watershed, such as the weathering of geological parent material. We presumed that hydrologic distance measures would better represent finer-grain processes related to flow connectivity. Dent and Grimm (1999) investigated the effect of flood events on spatial correlation in three chemical response variables using SHD and found that their range was affected by flood frequency. The range increased immediately following a flood event and then decreased over time, which created heterogeneous patterns of chemical concentration. The observed temporal pattern of correlation was similar for all response variables, but the range of their spatial correlation differed (Table I).

We found only one study where a WAHD measure was used to explore spatial patterns of correlation in water chemistry. Cressie and others (2005) developed an asymmetric hydrologic distance measure weighted by stream order, which was based on the work of Ver Hoef and others (2007). They considered a mixture of autocorrelation functions based on SLD and WAHD and determined that spatial

TABLE I
A summary of studies that have explored patterns of spatial autocorrelation in stream chemistry data using straight-line distance (SLD) and symmetric hydrologic distance (SHD)

Response variable	Geographic location	Distance measure[a]	Nugget	Sill	Range (km)	Autocorrelation function
Yuan, 2004						
Nitrate (NO^3)	Maryland,	SLD	N/A	N/A	49	Exponential
Sulfate (SO_4)	USA	SLD	N/A	N/A	68	Exponential
Kellum, 2003						
Acid nuetralizing capacity	Mid-Atlantic Highlands, USA	Anisotrophic SLD	0.104	0.21	160.93	Exponential
Gardner et al., 2003						
Temperature	Catskill	SLD	2.0	3.5	6	Spherical
Temperature	Mountains,	SHD	2.0	3.5	7.5	Spherical
Temperature	New York, USA	SHD weighted by stream order	0	5.2	10	Spherical
Dent and Grimm, 1999						
Nitrate-nitrogen						
2 week post flood	Central	SHD	N/A	N/A	>3	Spherical
2 month post flood	Arizona,	SHD	4.70	31.5	0.401	Spherical
9 month post flood	USA	SHD	195.00	2265	0.359	Spherical
Soluble Reactive Phosphorus						
2 week post flood		SHD	N/A	N/A	>3	Spherical
2 month post flood		SHD	N/A	N/A	>3	Spherical
9 month post flood		SHD	9.50	120	1.068	Spherical
Conductivity						
2 week post flood		SHD	N/A	N/A	>3	Spherical
2 month post flood		SHD	N/A	N/A	>3	Spherical
9 month post flood		SHD	1.00	973	1.025	Spherical
Cressie et al., 2005						
Dissolved Oxygen	Southeast Queensland, Australia	SLD & WAHD	0.88	1.043	6.07	Spherical

[a] Unless otherwise noted, covariance parameters for straight-line distance are isotropic.

dependence between instream dissolved oxygen measurements was better represented by SLD.

More than one distance measure can be used to explain variability in stream chemistry data (Gardner *et al.*, 2003), which is useful because we expect spatial correlation to differ between temporal and spatial filter scales (Levin, 1992; Poff, 1997). The distance measure and scale must be appropriate for the

ecological process being studied. However, applying geostatistical techniques to stream networks is a relatively new field of research and the limited findings to date do not clearly indicate which distance measure to use. To our knowledge, Cressie and others (2005) is the only study to quantitatively fit a variety of distance measures to one dataset to determine which best explains the variability in stream chemistry data. Gardner and others (2003) also compared predictions made using three distance measures, but used the spherical autocorrelation function to generate correlations based on hydrologic distances. The spherical autocorrelation function is invalid for pure hydrologic distance measures because the correlation matrices may contain negative eigenvalues, may produce negative variance estimates, and are not guaranteed to be positive definite (Ver Hoef *et al.*, 2007).

Our goal is to provide a detailed investigation into patterns of spatial correlation in eight chemical response variables collected throughout the state of Maryland, USA. We derive lumped watershed covariates, such as mean elevation or percent geology type in the watershed, using a geographical information system (GIS) and use them to explain the broad-scale trend in the mean of the stream chemistry data. The geostatistical models are based on SLD, SHD, and WAHD. We generate predictions and make a statistical comparison to determine which distance measure best explains the variability in each chemical response variable.

2. Methodology

2.1. DATA

The Maryland Biological Stream Survey (MBSS) data (Figure 2) were collected throughout Maryland by the Department of Natural Resources (DNR) (Mercurio *et al.*, 1999). Maryland is a geographically diverse state that can be divided into three general provinces: the Coastal Plain, the Piedmont, and the Appalachian Plateau (Boward *et al.*, 1999). The Coastal Plain borders Chesapeake Bay and produces low gradient streams with sandy gravel substrates. Elevation increases from east to west and Piedmont streams are characterized by steeper slopes and rock or bedrock substrates. The Appalachian Plateau is the westernmost region and is a diverse composition of valleys, sloping mountains, and steep ridges. Streams generally have rocky substrates, but range from low gradient meandering streams to steep cascading streams.

The Maryland DNR used a probability-based survey to collect chemical, physical, and biological data from first, second, and third order non-tidal streams in 17 interbasins throughout the state (Mercurio *et al.*, 1999). A stratified random sample was collected from each interbasin based on Strahler stream order. The number of samples collected per stream order was proportional to the number of stream order miles within the interbasin. Ten chemical variables were collected:

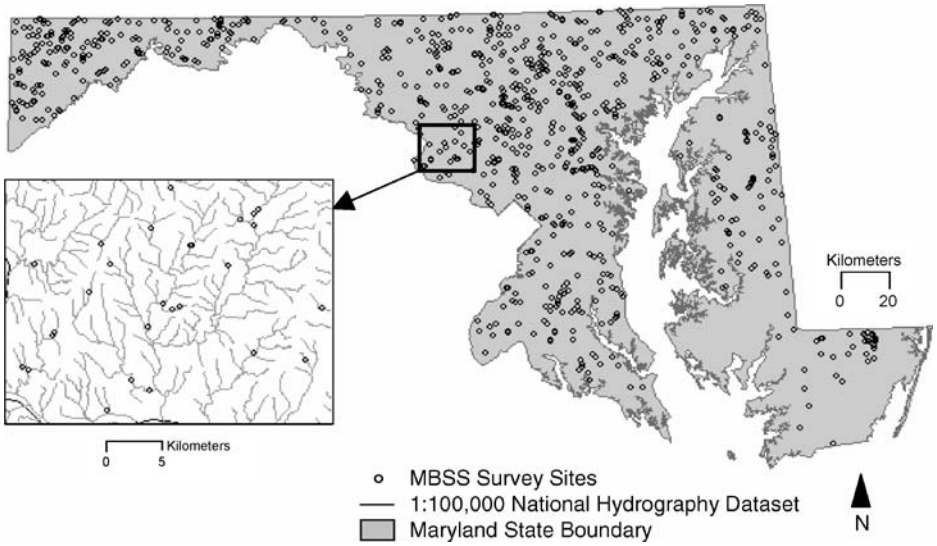

Figure 2. The Maryland Biological Stream Survey (MBSS) dataset includes chemical, physical, and biological data, which were collected throughout Maryland by the Maryland Department of Natural Resources during 1995, 1996, and 1997.

acid neutralizing capacity (ANC), in-situ conductivity, conductivity measured in the lab (CONDLAB), dissolved organic carbon (DOC), dissolved oxygen (DO), nitrate-nitrogen (NO^3), in-situ pH, pH measured in the lab (PHLAB), sulfate (SO_4), and temperature (TEMP). However, we did not include in-situ pH and conductivity in our analysis. In total, 955 sites were visited during 1995, 1996, and 1997.

The stream network and survey site coordinates were pre-processed in a GIS to ensure that sites were positioned on the correct stream segment. There are a variety of reasons why it is rare for GIS data collected within a stream to fall directly on a line segment representing a stream. Though the spatial accuracy of points collected using a global positioning system are becoming more precise, they still have some error (Bolstad *et al.*, 2005). Some stretches of river can move (e.g. meander) slightly from their mapped position (Dunne and Leopold, 1978). Streams are often represented by lines (USGS, 1999) and so samples collected on the banks of a large river may not fall directly on a line segment. Digital streams datasets may contain mapping errors and generalizations, such as the absence of small tributaries and the generalization of form, which are found when streams are represented at coarser scales (Veregin, 2000). As a result of these data problems, we discarded 74 sites because the survey stream could not be identified. In addition, there were a minimal number of missing data values for each chemical response variable. These sites were not completely eliminated from further analysis since more than one variable was collected at each site. Instead, they were temporarily removed from the analysis when a response variable contained no data.

Distance matrices were generated for SLD, SHD, and WAHD measures. We projected the data from latitude/longitude to Albers Equal Area projection (North American Datum 1983 based on the GRS1980 spheroid) before calculating the distance measurements. Projecting these data were necessary because distance between points calculated directly from latitude/longitude coordinates have a known, systematic bias associated with increasing latitude. The SLD matrix was calculated in R statistical software package (Ihaka and Gentleman, 1996) using the easting and northing values as x, y coordinates. The SHD, asymmetric hydrologic distance, and spatial weights matrices were calculated in a GIS using programs written in Visual Basic for Applications for ArcGIS version 8.3 (ESRI, 2002).

The spatial weights were used to develop the WAHD measure and represent the relative influence of one site on another. The weights were based on watershed area, which we use as a rough proxy for discharge volume. The spatial weights were generated by calculating the upstream watershed area for the downstream node of each segment in the stream network using a GIS. We defined a stream segment as the portion of a stream located between two confluences. When survey sites fell midway along a segment it was split into two separate stream segments. At each confluence or survey site in the network, the total upstream watershed area was calculated by summing the watershed area for the incoming stream segments. The proportional influence for each incoming segment was calculated by dividing its watershed area by the total upstream watershed area at the confluence or survey site. Every stream segment in the network contained its proportional influence on the segment directly downstream when this process was complete. Then, we located the path between flow connected sites and calculated the influence of one site on another, which was equal to the product of the segment proportional influences found in the path. The spatial weights matrix was simply an n by n matrix that contained the square root of the influence for all pairs of sites, which we used to maintain stationarity of the variances (Ver Hoef *et al.*, 2007). If two sites were not connected by flow the spatial weight was equal to zero and a sites influence on itself was equal to one. The GIS methods used to generate the hydrologic distance matrices and spatial weights matrix were lengthy, but are not the focal point of this manuscript (but see Peterson, 2005).

2.2. Statistical analysis

2.2.1. *Initial Covariate Selection*

The MBSS dataset contains lumped watershed attributes for each survey site (Mercurio *et al.*, 1999), which we used as potential covariates (Table II). In addition, we included the survey year, Level III Omernik's ecoregion (Omernik, 1987), mean elevation in the watershed, and geographic location. The variance inflation factor (VIF) collinearity statistic (Helsel and Hirsch, 1992) indicated that five potential covariates were significantly correlated with other covariates (VIF > 10) so we removed them from further analysis. We also created a validation set

TABLE II
Potential covariates

Covariate	Description
AREA	Watershed area (ha)
BARREN	% Barren
WATER	% Water
HIGHURB	% High intensity uban
LOWURB	% Low intensity urban
PASTUR	% Hay/pasture/grass
PROBCROP	% Probable row crop
ROWCROP	% Row crop
CONIFER	% Conifer or evergreen forest type
DECIDFOR	% Deciduous forest type
MIXEDFOR	% Mixed forest type
EMERGWET	% Emergent Wetlands
WOODYWET	% Woody wetlands
COALMINE	% Coalmine
EASTING	Easting - Albers Equal Area Conic
NORTHING	Northing - Albers Equal Area Conic
ELEV	Mean elevation in the watershed
YR96	Sample Year 1996
YR97	Sample Year 1997
ER67	Omernik's Level 3 Ecoregion 67
ER69	Omernik's Level 3 Ecoregion 69
ARGPERC	% Argillaceous rock type
CARPERC	% Carbonic rock type
FELPERC	% Felsic rock type
MAFPERC	% Mafic rock type
SILPERC	% Siliceous rock type

for each chemical response variable, which contained a unique set of 100 randomly selected sites (without replacement). These data were set aside in order to assess the accuracy of the final models.

We reduced the number of covariates due to the considerable processing time required for geostatistical model selection. For example, it was necessary to invert the matrices (881×881), fit the three covariance parameters, and fit the five regression coefficients at each model iteration. We acknowledge that setting a limit on the number of covariates may have resulted in the omission of additional covariates that possess significant predictive ability. However, our goal was to explain enough variability in the data to recognize and compare patterns of spatial autocorrelation in the chemical response variables. We do not believe that the possible omission of additional covariates would significantly affect the results of this study.

TABLE III
Summary statistics for chemical response variables

Response	Transformation	N	Min	1st Qu.	Median	Mean	3rd Qu.	Max
ANC (μeq/l)	log 10($x + 320$)	874	2.36	2.73	2.85	2.88	3.00	3.75
CONDLAB (μmho/cm)	log 10($x + 1$)	874	1.45	2.06	2.20	2.20	2.34	3.52
DOC (mg/l)	log 10($x + 1$)	877	0.00	0.40	0.49	0.56	0.70	1.53
DO (mg/l)	none	826	1.10	7.50	8.40	8.22	9.30	12.60
NO^3 (mg/l)	log 10($x + 1$)	873	0.00	0.22	0.42	0.44	0.64	1.07
PHLAB	none	866	4.40	6.79	7.16	7.12	7.45	8.90
SO_4 (mg/l)	none	870	0.32	0.93	1.10	1.10	1.24	2.72
TEMP (°C)	none	842	11.40	17.20	19.10	19.36	21.50	28.90
Covariate								
LOWURB	none	NA	0.00	0.10	0.97	6.70	5.05	76.74
HIGHURB	none	NA	0.00	0.00	0.12	1.12	0.63	22.90
WOODYWET	none	NA	0.00	0.00	0.80	2.91	3.13	34.96
PROBCROP	none	NA	0.00	10.34	19.15	20.27	29.05	85.47
ROWCROP	none	NA	0.00	2.51	6.72	9.46	13.36	50.44
PASTUR	none	NA	0.00	4.00	11.90	15.00	23.17	62.07
DECIDFOR	none	NA	1.74	16.52	26.05	32.66	45.06	98.43
CONIFER	none	NA	0.00	0.81	1.98	3.92	4.60	37.85
MIXEDFOR	none	NA	0.00	2.33	4.97	6.31	8.56	29.45
COALMINE	none	NA	0.00	0.00	0.00	0.18	0.00	12.91
WATER	none	NA	0.00	0.05	0.17	0.28	0.34	4.64
AREA (ha)	none	NA	11.63	385.46	1264.64	2602.13	3407.05	29068.57

*YR96, YR97, ER67, ER69, and NORTHING were also significant covariates, but were not included in the table since they are categorical data.

We used a Leaps and Bounds algorithm (Furnival and Wilson, 1974) to find the "best" set of five covariates, based on Mallow's Cp statistic (Neter *et al.*, 1996), for each chemical response variable and used them to develop a general linear model (GLM). We checked the model residuals for signs of non-normality and transformed ANC, CONDLAB, DOC, and NO^3 using a log10($x + n$) transformation (Table III). The data contained a limited number of extreme values that could not be explained by the models we proposed. If these data were included, they would have a disproportionate influence on the fitted autocorrelation function, and the analysis would be seriously compromised. Rather than use a formal robust estimation procedure (Rousseeuw and Leroy, 1987), we chose to proceed informally and simply removed the most extreme values. We used a hypothesis-testing paradigm to detect data values that produced a Studentized residual with a significance level less than 0.001.These data values were excluded from further analysis. We also calculated summary statistics for the response variables and the significant covariates.

2.2.2. *Model Parameter Estimation*

We restricted the model space to all possible linear models using the five explanatory variables determined by the initial covariate selection process described above. For a given distance measure there were 32 ($2^5 = 32$) competing models. Hence, four sets of 32 models were developed for each chemical response variable. The first set consisted of non-spatial models developed using GLM as a baseline to determine whether geostatistical models provided additional predictive ability. Each of the remaining sets assumed a geostatistical model using one of three distance measures: SLD, SHD, or WAHD. We also assumed that the model residuals were normally distributed with mean zero and variance-covariance matrix $\Sigma = \sigma^2\Omega$, where σ^2 is the variance and $\Omega = \Omega(d;\theta)$ is the correlation matrix. Note that Ω is a function of the distance between sites, d, given the autocorrelation parameter vector, θ. Therefore, the model for response variable Z is written in matrix notation as $Z = X\beta + \varepsilon$ where $\varepsilon \sim N(0, \sigma^2\Omega)$. Here X is the $n \times p$ design matrix of covariates, β is a vector of coefficients of length p, and ε is a vector of n (correlated) errors.

The log-likelihood function of the parameters $(\theta, \beta, \sigma^2)$ given the observed data, Z, is

$$\ell(\theta, \beta, \sigma^2; Z) = -\frac{n}{2}\log(2\pi) - \frac{1}{2}\log|\sigma^2\Omega| - \frac{1}{2\sigma^2}(Z - X\beta)'\Omega^{-1}(Z - X\beta). \quad (2)$$

Maximizing the log-likelihood (2) with respect to β and σ^2 yields $\hat{\beta} = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}Z$ and $\hat{\sigma}^2 = \frac{(Z-X\hat{\beta})'\Omega^{-1}(Z-X\hat{\beta})}{n}$. Both maximum likelihood estimators (MLE) can be written as functions of θ alone. Thus, we derive the *profile* log-likelihood function (Cressie, 1993) by substituting the MLEs back into (2)

$$\ell_{\text{profile}}(\theta; \hat{\beta}, \hat{\sigma}^2, Z) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\hat{\sigma}^2) - \frac{1}{2}\log|\Omega| - \frac{n}{2}. \quad (3)$$

The primary advantage to using the profile log-likelihood is that it reduces the dimensionality of the problem, which can reduce the amount of time required to find a numerical solution. This is especially important when there are a large number of models to compare.

The correlation matrix, Ω, is computed using the exponential autocorrelation function defined as

$$C_1(d;\theta_1,\theta_2) = \begin{cases} 1 & \text{if } d = 0 \\ (1-\theta_1)\exp(-d/\theta_2) & \text{if } d > 0', \end{cases} \quad (4)$$

where θ_1 is the *proportion* of nugget effect and θ_2 is the range parameter. The nugget is estimated by $\hat{\theta}_1\hat{\sigma}^2$ where θ_1 is restricted between zero and one. The approximate

range over which sites are considered to be correlated is $3\theta_2$ (Cressie, 1993). The value d represents the distance between any two sites relative to the distance measure, e.g., SLD or SHD. The correlation matrix for the WAHD measure is generated by taking the Hadamard (element-wise) product of (4) and the spatial weights matrix (Ver Hoef *et al.*, 2007). We restricted our analysis to the exponential autocorrelation function because it is currently the single known valid autocorrelation function for SHD (Ver Hoef *et al.*, 2007). We assumed that the errors were independent for the GLM model, i.e., Ω can be replaced with the identity matrix, I_n.

The MLE for θ is found by maximizing the profile log-likelihood (3) using a quasi-Newton method (Byrd *et al.*, 1995), which is in turn used to compute the MLEs for the model parameters: β and σ^2. To promote numerical stability, we standardized the response and explanatory variables to have mean zero and unit variance and scaled the distances to fall between zero and one.

2.2.3. *Model Selection and Model Performance*

We used the spatial AICC statistic (Hoeting *et al.*, 2006) to select the "best" GLM and the three "best" geostatistical models for each response variable: one for each of the three distance measures. The spatial AICC statistic is defined as

$$\text{AICC} = -2\ell_{\text{profile}}(\theta; \beta, \sigma^2, Z) + 2n\frac{p+k+1}{n-p-k-2}, \tag{5}$$

where n is the number of observations, $p-1$ is the number of covariates, and k is the number of autocorrelation parameters. To select the "best" GLM model we set $k = 0$. The parameter k was set to two for the remaining three distance measures.

The model with the smallest AICC from each set was used to generate predictions using the universal kriging algorithm (Cressie, 1993). We used a split-sample approach to calculate the mean square prediction error (MSPE) for each model using validation sets that were set aside at the beginning of the analysis. The MSPE is defined as

$$\text{MSPE} = \frac{\sum_{i=1}^{n_p}(Z_i - \hat{Z}_i)^2}{n_p}, \tag{6}$$

where Z_i is the observed value at site i, $\hat{Z}_i$ is the predicted value at site i, and n_p is the total number of predictions. The MSPE was computed using a unique validation set for each response variable, but the four models within a response variable were tested using the same validation set. Models with small MSPE are desirable. The MSPE provided a way to compare models constructed using different distance measures and to determine which measure, if any, was more able to account for the variability in the response variable. In addition, we calculated the squared Pearson correlation coefficient (r^2) between the predictions and observations.

3. Results

Summary statistics for the chemical response variable and significant covariate distributions are provided in Table III.

The spatial neighborhood produced by each distance measure differs, which affects the number of neighboring sites, as well as, the median, mean, and maximum separation distance between sites (Table IV). Asymmetric hydrologic distance had considerably fewer pairs of neighboring sites compared to SLD and SHD. The minimum separation distance between neighboring sites was similar for all distance measures, but the asymmetric hydrologic distance measure had a shorter median, mean, and maximum value than the other distance measures. SHD consistently had the largest median, mean, and maximum separation distance.

We used five covariates in the model selection process (Table V) and our results show that the models with the lowest spatial AICC value tended to be complex, meaning that they included a large number of covariates (Table VI). Covariates

TABLE IV
Summary statistics for distance measures in kilometers for dissolved oxygen ($n = 826$)

Distance measure	N Pairs	Min	Median	Mean	Max
Straight-line distance	340725	0.05	101.02	118.16	385.53
Symmetric hydrologic distance	62625	0.05	156.29	187.10	611.74
Pure asymmetric[a] hydrologic distance	1117	0.05	4.49	5.83	27.44

[a]Asymmetric hydrologic distance was not weighted in the summary statistics.

TABLE V
Five significant covariates for each chemical response variable selected using a leaps and bounds regression

Response	Significant covariates
ANC (μeq/l)	PASTUR, LOWURB, WOODYWET, YR96, YR97
CONDLAB (μmho/cm)	HIGHURB, LOWURB, COALMINE, YR96, NORTHING
DOC (mg/l)	WOODYWET, CONIFER, MIXEDFOR, LOWURB, NORTHING
DO (mg/l)	DECIDFOR, HIGHURB, WOODYWET, YR96, YR97
NO_3 (mg/l)	PASTUR, PROBCROP, ROWCROP, LOWURB, WATER
PHLAB	PROBCROP, DECIDFOR, WOODYWET, AREA, CONIFER
SO_4 (mg/l)	LOWURB, COALMINE, NORTHING, ER67, ER69
TEMP ($^\circ$C)	PROBCROP, LOWURB, WATER, YR96, YR97

TABLE VI

Autocorrelation parameter estimates, mean square prediction error (MSPE), and squared Pearson correlation coefficient (r^2) for the general linear model (GLM), straight-line distance (SLD), symmetric hydrologic distance (SHD), and weighted asymmetric hydrologic distance (WAHD) model for each unscaled response variable. The MSPE and r^2 values were calculated using the observed and predicted values contained in the validation set

Response variable	Distance measure	Model	MSPE	r^2	Nugget%	Sill	Range (km)
ANC	GLM	32	293792.34	0.411	NA	NA	NA
	SLD	32	50672.47	0.899	1.90	0.388	26.85
	SHD	32	55861.61	0.877	2.00	0.286	57.57
	WAHD	32	87203.41	0.843	0.90	0.644	47.76
CONDLAB	GLM	32	34969.41	0.712	NA	NA	NA
	SLD	32	11623.06	0.921	0.90	0.961	12.47
	SHD	32	3239.63	0.959	1.30	0.573	27.67
	WAHD	32	4014.98	0.948	1.10	0.569	45.27
DOC	GLM	32	7.85	0.520	NA	NA	NA
	SLD	32	5.46	0.644	1.54	0.282	56.39
	SHD	32	5.37	0.656	2.89	0.693	180.79
	WAHD	32	5.47	0.649	1.99	0.734	82.32
DO	GLM	32	1.91	0.294	NA	NA	NA
	SLD	31	1.58	0.414	5.45	0.202	62.77
	SHD	31	1.64	0.392	7.04	0.283	301.76
	WAHD	32	1.74	0.355	3.98	0.263	82.32
NO^3	GLM	32	1.14	0.671	NA	NA	NA
	SLD	32	0.82	0.772	3.60	0.593	20.78
	SHD	32	0.75	0.783	7.40	0.957	45.13
	WAHD	32	0.95	0.725	6.50	0.937	73.30
PHLAB	GLM	32	0.16	0.504	NA	NA	NA
	SLD	31	0.11	0.663	4.70	0.647	16.35
	SHD	31	0.10	0.679	6.40	0.500	36.46
	WAHD	32	0.11	0.663	3.50	0.503	33.99
SO_4	GLM	32	363.78	0.190	NA	NA	NA
	SLD	20	210.14	0.400	1.81	0.271	23.46
	SHD	28	259.42	0.360	3.06	0.443	40.84
	WAHD	28	292.21	0.286	1.76	0.922	82.32
TEMP	GLM	32	8.81	0.177	NA	NA	NA
	SLD	32	7.72	0.278	1.25	0.310	6.90
	SHD	32	7.49	0.298	4.20	0.702	14.03
	WAHD	32	7.37	0.309	1.88	0.473	15.49

were systematically added during model selection so that model 1 represents the null model (no covariates) and model 32 the full model (all five covariates). The full model was selected for every GLM and for 17 of the 24 geostatistical models (Table VI).

The models for DO, PHLAB, and SO_4 displayed differences in complexity. Model 31 was selected for the DO and PHLAB models based on SLD and SHD. The DO model included all of the significant covariates except DECIDFOR, while the PHLAB model excluded PROBCROP (Table V). This was not surprising because exploratory data analysis indicated that DO and PHLAB were weakly correlated with DECIDFOR and PROBCROP ($r^2 = 0.05$ and $r^2 = 0.01$, respectively). However, EASTING explained 44% of the variability in PROBCROP and 63% of the variability in DECIDFOR. The simplification of the model suggests that the correlations produced using SLD and SHD successfully represented the relatively small amount of variability in the response that was also explained by DECIDFOR and PROBCROP. In contrast, the full model was selected for the WAHD. Although models 31 and 32 are similar, the difference in model complexity indicates that the correlations based on WAHD contained less information about spatial correlation in the response variable compared to those produced using the SLD and SHD measures. SO_4 was the only chemical response variable where a simpler model was consistently selected for all distance measures (Table VI). NORTHING was omitted from the models based on SHD and WAHD and ER67 was omitted from the SLD model (Table V). We believe that NORTHING and ER67 also represent broad-scale trends on the landscape related to spatial location.

The nugget estimates produced using ML differed between response variables. ANC, CONDLAB, NO^3, and PHLAB models had small nugget estimates, which included less than 7.5% of the variability in the data (Table VI). The nugget estimates for the DOC, SO_4, and TEMP models were larger and represented between 12.5% and 42% of the variability. Nugget estimates for DO were noticeably larger for all three distance measures compared to other chemical response variables and included between 39% and 70.4% of the variability in the data.

The ML estimates for the range parameter varied greatly with respect to distance measure and response variable (Table VI). We found that SLD produced the shortest range for every chemical response variable. The mean range for SLD, SHD, and WAHD were 28.2 km, 88.03 km, and 57.8 km, respectively. TEMP consistently produced the shortest range values and DO generated the largest range values for every distance measure. It should be noted that the ML range parameter estimates for the DOC, DO, and SO_4 models based on WAHD fell at the (user-defined) upper bounds of the optimization method.

The geostatistical models consistently explained more variability in the chemical response variables than the GLM (Table VI). The most dramatic differences were found in ANC, CONDLAB, and SO_4 where the r^2 value increased by 49, 24, and 21 percentage points, respectively. The geostatistical models for ANC, CONDLAB,

DOC, NO^3, and PHLAB models generated predictions that had a strong correlation to the observed values ($r^2 = 0.63$ to 0.95). However, the correlation between the geostatistical model predictions generated by the DO, SO_4, and TEMP models and the observations was weaker ($r^2 = 0.28$ to 0.41).

A comparison of the MSPE values suggests that more than one distance measure performed well for most chemical response variables (Table VI). The MSPE values for the DOC, TEMP, and PHLAB models did not exhibit clear differences relative to distance measure. SLD and SHD models developed using NO^3, ANC, and DO had similar MSPE values and produced more accurate predictions than the equivalent models using WAHD. In contrast, the SHD and WAHD models provided more accurate CONDLAB predictions than the SLD model. SO_4 was the only chemical response variable that displayed obvious differences between the models using different distance measures. The SLD model had the smallest MSPE value and produced more accurate predictions compared to the other distance measures.

4. Discussion

The summary statistics for the distance measures (Table IV) demonstrate how a distance measure can significantly influence the way that spatial relationships are represented in a stream network. Not only does it affect the distance between neighbors and their relative influence, but it also dictates the form and size of the spatial neighborhood. For example, a different set of neighboring sites may be included in the spatial neighborhood when SLD, SHD, and WAHD distance measures are used. The geostatistical model estimates the local deviation from the mean in the data at unobserved sites based on the proximity and value of observed neighboring sites. When ecologically unrelated sites are included in the spatial neighborhood the estimates are based on incorrect information, which may negatively impact the accuracy of the geostatistical model predictions. Therefore, decisions concerning which distance measure to use are not just statistical choices, but should also be founded on the specific characteristics of the ecological process of interest and the research questions.

It was difficult to locate and obtain a dataset that was suitable for a regional analysis of spatial correlation in stream networks. We examined over 35 Environmental Protection Agency (EPA) datasets and most had large minimum separation distances between sites or few neighboring sites when a WAHD measure was used (e.g. Colorado Regional Environmental Mapping and Assessment Program dataset (USEPA, 1993)). The EPA uses a probability-based random survey design based on stream order to estimate regional stream conditions (Herlihy *et al.*, 2000). This method is useful because it provides a statistical inference about the entire population of streams, within stream order, over a large area. However, it was designed to maximize spatial independence of survey sites, and consequently does not adequately represent spatial relationships in stream networks based on hydrologic distance measures.

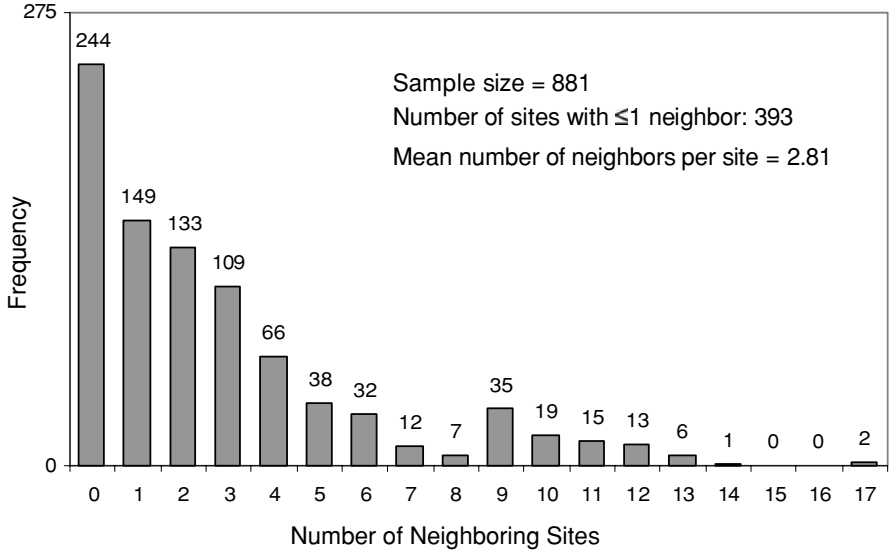

Figure 3. Summary of spatial relationships between neighboring sample sites for acid neutralizing capacity using the weighted asymmetric hydrologic distance measure.

The MBSS data were also collected using a probability-based design (Mercurio *et al.*, 1999), but the sampling density was greater than other water chemistry datasets. Despite our large sample size ($n = 881$), 244 sites did not have neighbors (Figure 3). The validation sites were randomly selected and it is likely that a considerable number did not have neighbors using the WAHD.

The models based on WAHD consistently produced more accurate predictions than the GLM models (Table VI). However, when a spatial neighborhood was deficient or absent for a specific site, the WAHD model performed in a manner similar to the GLM (Figure 4). This is a common feature for geostatistical models. The associated standard error for prediction sites with many observed neighbors is small compared to sites that have few (or no) neighbors. Thus, the WAHD model had the ability to explain the broad-scale mean in the data, but did not provide additional predictive ability at that site. The WAHD model generally explained more variability in the data as the number of neighboring sites increased (Figure 4). However, notable exceptions occurred when a site had neighbors with similar watershed conditions, but significantly different water chemistry values. Although the WAHD models were comparable to the SLD and SHD models, we believe that their performance may have been hindered by the survey design used to collect the raw data and the consequent lack of neighboring sites.

The GLM predictions also improved as the number of neighbors increased (Figure 4) because clusters of sites in space tend to have similar covariate values, i.e., they are positively correlated. Even though spatial location was not included in the model, the statistical regression was pulled towards the cluster of similar

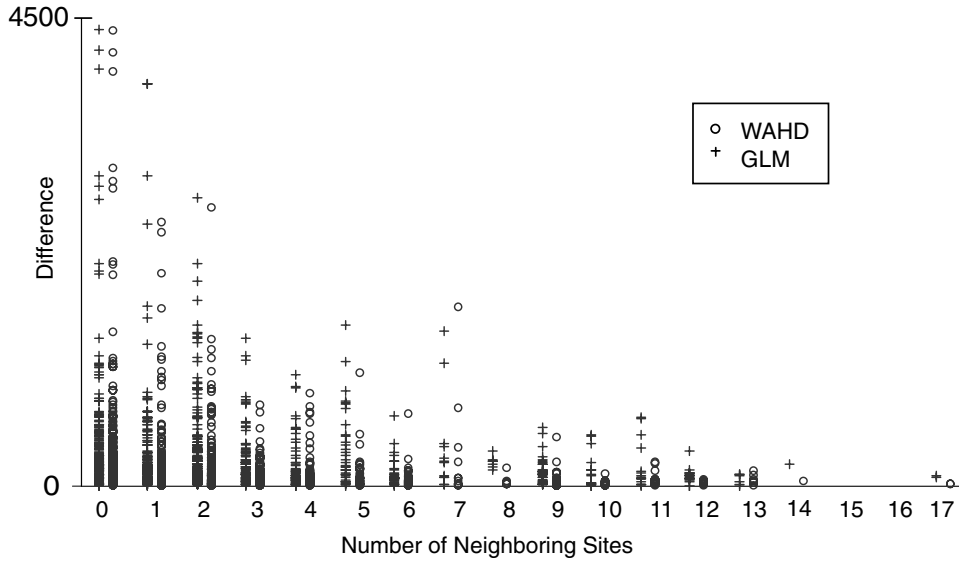

Figure 4. Difference between observed values and general linear model (GLM) predictions or observed values and weighted asymmetric hydrologic distance (WAHD) predictions according to the number of neighbors.

values. Thus, the GLM contained hidden spatial information, which allowed the non-spatial model to explain additional variability in the data when a greater number of neighbors were present.

Our results provide evidence that patterns of spatial correlation exist in stream water chemistry. The predictive ability of every geostatistical model was greater than the GLM (Table VI) and in most cases more than one distance measure could be used to quantify the additional variability in stream chemistry data. SLD clearly does not represent the flow connectivity between sites in a stream network. SHD is similar in this respect because sites need not be connected by flow to be neighbors. The SHD separation distances and range values are consistently larger than those produced using SLD (Table IV). This is intuitive since the distance traveled between two sites increases when movement is restricted to the sinuous network. If the models based on SLD had performed poorly, we could assume that water chemistry was dominated by instream processes at a regional scale. However, the SLD models were never substantially inferior, which leads us to believe that the SLD, SHD, and WAHD measures are representing patterns of spatial correlation in continuous coarse-scale variables, such as geology type, that influence stream chemistry rather than the movement of chemicals through a stream network.

The chemical response variables that produced models with the greatest amount of predictive ability tend to be strongly related to coarse-scale landscape variables. ANC, PHLAB, NO^3, and CONDLAB are all significantly related to landscape

variables that are not restricted to watershed boundaries, such as geology type (Kellum, 2003), agricultural and urban areas (Herlihy *et al.*, 1990; Gray, 2004), and the atmospheric deposition of nitric and sulfuric acids (Angelier, 2003). The small nuggets estimated by the geostatistical models indicate that the survey scale was fine enough to capture the majority of variability in the data.

The range values differ between ANC, NO^3, PHLAB, and CONDLAB and we believe that they represent the coarse scale of heterogeneity in the ecological processes that control them. However, the range parameter is estimated by modeling the spatial structure in the residual error and so it does not provide information about which ecosystem processes are producing patterns of spatial correlation. For example, unexplained variability may be related to a strongly influential model covariate that was not proposed during model selection. In contrast, the influential ecological process may be related to a significant model covariate, but information may have been lost by using coarse-scale lumped (non-spatial) watershed covariates. Therefore, what follows is simply knowledgeable speculation about some, but not all, potential sources of spatial correlation in the data.

ANC and NO^3 models contained coarse-scale agricultural and urban covariates and produced similar range values. Thus, runoff or leaching from agricultural and urban areas may strongly influence ANC and NO^3 concentrations in the stream. The PHLAB model contained an agricultural covariate, but the range estimates were slightly shorter. PHLAB was also correlated to forest and wetland watershed area, which would be expected to exhibit more heterogeneous patterns on the landscape. CONDLAB produced the shortest range values of the models with strong predictive ability and was affected by urban land uses. However, percent coal mine area in the watershed was another influential covariate, which could actually be considered a point source of pollution. Coal mines are generally located in the western portion of Maryland, but would not produce a continuous pattern on the landscape. For example, the presence of one coal mine does not indicate that another coal mine will be located in a neighboring watershed. The relatively unpredictable distribution of coal mines may explain the short range values for CONDLAB. It is also not surprising that the models using SHD and WAHD performed slightly better than the SLD model (Table VI) since the effects of coal mines are found downstream rather than in the adjacent watershed. Although the magnitude of the range values differ for ANC, NO^3, PHLAB and CONDLAB, they are all strongly influenced by the coarse-scale condition on the landscape, which reduces the effect of in stream processes at this survey scale.

The predictive ability of our models suggests that DOC is influenced by ecological processes that produce coarse-scale patterns of spatial correlation on the landscape. This is not surprising since the majority of stream DOC comes from allochthonous sources of organic matter such as dead terrestrial plant material, soil, groundwater inputs, and wetlands (Wetzel, 1992). In addition, DOC is transported from the watershed to the stream via overland, sub-surface, or base flow and the flow path of water affects the DOC concentration of stream water (Qualls and

Haines, 1992; Mulholland, 2003). The ML range estimate for the WAHD model was set to the upper limit in the optimization, but we feel that this is reasonable because the range values for SLD and SHD were also quite large. Interestingly, models generated to predict DOC were the only models with considerable predictive ability and relatively large nugget estimates. The large nugget estimate may result from our failure to represent the terrestrial water flow in the model, point sources of organic pollution, or fine scale instream processes that significantly affect stream DOC.

SO_4 is also influenced by coarse-scale processes, such as atmospheric inputs of sulfur and the weathering of sulfate minerals, and finer-scale processes related to the mineralization of organic sulfur and the adsorption and desorption of sulfate (e.g. Alewell *et al.*, 1999). However, our results were inconclusive and were similar to other studies that have failed to establish strong correlations between SO_4 and watershed landcover categories (Herlihy *et al.*, 1998).

The SO_4 model using SLD had the smallest MSPE, but still had little predictive capability. All models produced large nugget values and the range value for SO_4 based on WAHD was set at the maximum separation distance. These results could indicate that finer-scale processes dominate patterns of spatial correlation in SO_4 and that coarse-scale processes are inconsequential. However, the ecological literature does not support that conclusion since the effects of atmospheric deposition of sulfuric acid on stream water chemistry have been well documented (Driscoll *et al.*, 2001). A more plausible explanation is that the model fails to represent a SO_4 input that does not produce a consistent spatial pattern on the landscape at the survey scale. The unexplained variability in SO_4 would significantly impair the predictive ability of the model and would essentially make coarse and fine scale patterns of spatial correlation indiscernible. The magnitude of the unexplained variability in the data suggests that more work is required to identify a better model.

The geostatistical models produced for TEMP and DO also had relatively little predictive ability. We concede that we are unable to conclusively determine the source of the residual error in these models, but we speculate that TEMP and DO are spatially and temporally variable over short distances (Hynes, 1960; Biggs *et al.*, 1990). They are somewhat influenced by coarse-scale factors, such as air temperature or nutrient inputs (Smith, 1981), but are most likely dominated by instream processes related to the biological oxygen demand or water depth. In addition, diurnal fluctuations in TEMP and DO resulting from changes in climate and the biological oxygen demand in the water may have contributed to the variability in the data. TEMP consistently produced the smallest range values and the nugget estimates were large, which provided further evidence that patterns of spatial correlation were occurring at a fine scale. All three distance measures performed equally well for TEMP, which again lead us to believe that the WAHD measure represented coarse-scale patterns of spatial correlation across the landscape. Although the geostatistical models for DO showed a slight improvement over the GLM model in the accuracy of predictions, the model fit was extremely poor. The nugget estimates for DO were consistently large, which indicated that there was

little to no spatial correlation between sites. The range values were also unrealistically large (Table VI) given the ecological behavior of DO, which provided further evidence that the form of the selected model was inappropriate. Given the large nugget estimates and the poor predictive ability of the models, we believe that the dominant processes controlling TEMP and DO occur at a scale finer than our minimum separation distance.

It is difficult to compare the results of our study to those of other studies because there are regional differences in the ecological processes that affect water chemistry (Hill *et al.*, 2000). However, Yuan (2004) also used the MBSS dataset and found strikingly different range estimates for NO^3 and SO_4 based on SLD (Tables I and VI). We suspect that the differences result from the methods that were used to fit the autocorrelation functions. It appears that Yuan (2004) used a weighted least squares method to fit the autocorrelation function to the mean semivariance at each separation distance. Maximizing the likelihood function prevents inconsistencies in the autocorrelation parameters related to (empirical) semivariogram bin size. The methodology proposed by Hoeting and others (2006) allows for simultaneous fitting of the model and the error process. In addition, it can be used to compare different models using the spatial AICC.

5. Conclusions

The results of our study clearly demonstrate that spatial correlation exists in stream chemistry data at a relatively coarse scale and that geostatistical models improve the accuracy of predictions. The ranges and patterns of spatial correlation differ between chemical response variables, which are influenced by ecological processes acting at different spatial and temporal filter scales. We believe that the coarse-scale patterns we identified reflect the effects of unknown regional scale ecological processes that are better described using SLD. Yet, spatial patterns are likely to change with the grain of the survey scale and the configuration of survey sites. Therefore, we inevitably impose bias related to the minimum separation distance and spatial neighborhood (Levin, 1992). For example, coarse-scale patterns may not be easily discernable if the extent of our survey area were smaller. In addition, a dataset with shorter separation distances and a more dense spatial arrangement would likely reveal spatial patterns related to finer-scale processes. Our results provide information about the ability of SLD, SHD, and WAHD measures to account for additional variability in stream water chemistry at a regional scale. Further research is needed to assess the capacity of these distance measures to explain additional variability at finer scales, such as at the watershed or reach scale.

Agencies have invested substantial resources collecting datasets using probability-based random survey designs and our study demonstrates that these data can be used to predict water quality conditions throughout large areas. At present, SLD appears to be the most suitable distance measure for regional geostatistical

modeling of CONDLAB, ANC, NO^3, SO_4, PHLAB, DO, DOC, and TEMP in Maryland when the data are collected using a probability-based random survey design. SLD matrices are simple to calculate in statistical or GIS software, while the pre-processing time for hydrologic distance measures is greater. To fully explore the possible advantages of hydrologic distances, easy-to-use tools need to be widely available. At present, a regional geostatistical model based on SLD could be used by managers to predict water quality conditions for every stream segment within a large area. This methodology could potentially help states and tribes identify water quality impaired stream segments, which would allow agencies to focus additional field sampling efforts on potentially impaired sites.

It is likely that water chemistry contains both coarse and fine scales of spatial correlation and they may be better quantified using different distance measures. It is possible to describe multiple spatial patterns using different distance measures and to incorporate them into one geostatistical model (Cressie *et al.*, 2005). However, it is doubtful that datasets collected using a probability-based random survey design will be suitable for these types of models. Multi-scale geostatistical models may require new survey designs that more fully capture spatial correlation at a variety of scales using multiple distance measures. Survey designs that explain variability at multiple scales have been developed for terrestrial ecosystems (Shmida, 1984; Stohlgren *et al.*, 1995) and it should also be possible to develop them for stream networks.

Acknowledgements

We would like to thank Brian Bledsoe, Will Clements, Melinda Laituri, and the anonymous reviewers who provided useful comments on earlier drafts of this paper. The work reported here was generously supported by a STAR Research Assistance Agreement (CR-829095) awarded by the U.S. EPA to the Space Time Aquatic Resource Modeling and Analysis Program (STARMAP) at Colorado State University. However, this paper has not been formally reviewed by the EPA and the views expressed here are solely those of the authors. The EPA does not endorse any products or commercial services mentioned in this paper.

References

Akaike, H.: 1973, 'Information theory and an extension of the maximum likelihood principle', in: *Second International Symposium on Information Theory*, Budapest, Hungary, pp. 267–281.

Alewell, C., Mitchell, M. J., Likens, G. E. and Krouse, H. R.: 1999, 'Sources of stream sulfate at the Hubbard Brook Experimental Forest: Long-term analyses using stable isotopes', *Biogeochemistry* **44**, 281–299.

Altunkaynak, A., Ozger, M. and Sen, Z.: 2003, 'Triple diagram model of level fluctuations in Lake Van, Turkey', *Hydrology and Earth System Sciences* **7**, 235–244.

Angelier, E.: 2003, *Ecology of Streams and Rivers*, Science Publishers, Inc, Enfield, NH, USA.

Benda, L., Poff, N. L., Miller, D., Dunne, T., Reeves, G., Pess, G. and Pollock, M.: 2004, 'The Network Dynamics Hypothesis: How channel networks structure riverine habitats', *BioScience* **54**, 413–427.

Biggs, B. J., Jowett, I. J., Quinn, J. M., Hickey, C. W., Davies-Colley, R. J. and Close, M.: 1990, 'Ecological characterisation, classification, and modeling of New Zealand rivers: An introduction and synthesis', *New Zealand Journal of Marine and Freshwater Research* **24**, 277–304.

Bolstad, P., Jenks, A., Berkin, J. and Horne, K.: 2005, 'A comparison of autonomous, WAAS, real-time, and post-processed global positioning systems (GPS) accuracies in northern forests', *Northern Journal of Applied Forestry* **22**, 5–11.

Boward, D., Kayzak, P., Stranko, S., Hurd, M. and Prochaska, A.: 1999, *From the Mountains to the Sea: The State of Maryland's Freshwater Streams*, EPA/903/R-99/023, Maryland Department of Natural Resources, Monitoring and Non-tidal Assessment Division, Annapolis, MD, USA.

Byrd, R. H., Lu, P., Nocedal, J. I. , Zhu, C. and Siam, J.: 1995, 'A limited memory algorithm for bound constrained optimization', *Scientific Computing* **16**, 1190–1208.

Carroll, R. J. and Ruppert, D.: 1982, 'A comparison between maximum likelihood and generalized least squares in a heteroscedastic linear model', *Journal of the American Statistical Association* **77**, 878–882.

Chambers, P. A., Prepas, E. E. and Gibson, K.: 1992, 'Temporal and spatial dynamics in riverbed chemistry-The influence of flow and sediment composition', *Canadian Journal of Fisheries and Aquatic Sciences* **49**, 2128–2140.

Closs, G., Downes, B. and Boulton A.: 2004, *Freshwater Ecology*, Blackwell Science Ltd, Malden, MA, USA.

Colyer, W. T., Kershner, J. L. and Hilderbrand, R. H.: 2005, 'Movements of fluvial Bonneville cutthroat trout in the Thomas Fork of the Bear River, Idaho-Wyoming', *North American Journal of Fisheries Management* **25**, 954–963.

Cooper, S. D., Barmuta, L., Sarnelle, O., Kratz, K. and Diehl, S.: 1997, 'Quantifying spatial heterogeneity in streams', *Journal of the North American Benthological Society* **16**, 174–188.

Cressie, N.: 1993, *Statistics for Spatial Data Revised Edition*, John Wiley and Sons, New York, NY, USA.

Cressie, N., Frey, J., Harch, B. and Smith, M.: 2005, 'Spatial prediction on a river network', *Technical Report*, No. 747, The Ohio State University, Columbus, OH, USA.

Cumming, G. S.: 2002, 'Habitat shape, species invasions, and reserve design: Insights from simple models', *Conservation Ecology*, **6**, 3. [online] URL: http://www.consecol.org/vol6/iss1/art3/.

Davies, N. M., Norris, R. H. and Thoms, M. C.: 2000, 'Prediction and assessment of local stream habitat features using large-scale catchment characteristics', *Freshwater Biology* **45**, 343–369.

Dawson, J. J. C., Billett, M. F. and Hope, D.: 2001, 'Diurnal variations in the carbon chemistry of two acidic peatland streams in north-east Scotland', *Freshwater Biology* **46**, 1309–1322.

Dent, C. L. and Grimm, N. B.: 1999, 'Spatial heterogeneity of stream water nutrient concentrations over successional time', *Ecology* **80**, 2283–2298.

Driscoll, C. T., Lawrence, G. B., Bulger, A. J., Butler, T. J., Cronan, C. S., Eagar, C., Lambert, K. F., Likens, G. E., Stoddard, J. L. and Weathers, K. C.: 2001, *Acid Rain Revisited: Advances in Scientific Understanding Since the Passage of the 1970 and 1990 Clean Air Act Amendments*. Science Links™ Publication, Hubbard Brook Research Foundation.

Dunne, T. and Leopold, L. B.: 1978, *Water in Environmental Planning*, W. H. Freeman and Co., San Francisco, CA, USA.

Environmental Systems Research Institute (ESRI): 2002, *ArcGIS version 8.3*, Redlands, CA.

Frissell, C. A., Liss, W. J., Warren, C. E. and Hurley, M. D.: 1986, 'A hierarchical framework for stream habitat classification: Viewing streams in a watershed context', *Environmental Management* **10**, 199–214.

Furnival, G. and Wilson, R.: 1974, 'Regression by leaps and bounds', *Technometrics* **16**, 499–511.
Ganio, L. M., Torgersen, C. E. and Gresswell, R. E.: 2005, 'A geostatistical approach for describing spatial pattern in stream networks', *Frontiers in Ecology and the Environment* **3**, 138–144.
Gardner, B., Sullivan, P. J. and Lembo, A. J.: 2003, 'Predicting stream temperatures: Geostatistical model comparison using alternative distance metrics', *Canadian Journal of Aquatic Science* **60**, 344–351.
Gray, L.: 2004, 'Changes in water quality and macroinvertebrate communities resulting from urban stormflows in the Provo River, Utah, U.S.A.', *Hydrobiologia* **518**, 33–46.
Helsel, D. R. and Hirsch, R. M.: 1992, *Statistical Methods in Water Resources*, Elsevier Science Publishing Co., New York, NY, USA.
Herlihy, A. T., Kaufmann, P. R. and Mitch, M. E.: 1990, 'Regional estimates of acid mine drainage impact on streams in the Mid-Atlantic and southeastern United States', *Water, Air, and Soil Pollution* **50**, 91–107.
Herlihy, A. T., Stoddard, J. L. and Johnson, C. B.: 1998, 'The relationship between stream chemistry and watershed land cover data in the Mid-Atlantic region, U.S.A.', *Water, Air, and Soil Pollution* **105**, 377–386.
Herlihy, A. T., Larsen, D. P., Paulsen, S. G., Urquhart, N. S. and Rosenbaum, B. J.: 2000, 'Designing a spatially balanced, randomized site selection process for regional stream surveys: The EMAP Mid-Atlantic Pilot Study', *Environmental Monitoring and Assessment* **63**, 95–113.
Hill, B. H., Hall, R. K., Husby, P., Herlihy, A. T. and Dunne, M.: 2000, 'Interregional comparisons of sediment microbial respiration in streams', *Freshwater Biology* **44**, 213–222.
Hoeting, J. A., Davis, R. A., Merton, A. A. and Thompson, S. E.: 2006, 'Model selection for geostatistical models', *Ecological Applications* **16**, to appear.
Hynes, H. B.: 1960, *The Biology of Polluted Waters*, Liverpool University Press, Liverpool, England.
Ihaka, R. and Gentleman, R.: 1996, 'R: A language for data analysis and graphics', *Journal of the Computational and Graphical Statistics* **5**, 299–314.
Jobson, J. D. and Fuller, W. A.: 1980, 'A comparison between maximum likelihood and generalized least squares in a heteroscedastic linear model', *Journal of the American Statistical Association* **75**, 176–181.
Kellum, B.: 2003, 'Analysis and modeling of acid neutralizing capacity in the Mid-Atlantic Highlands area', Master's Thesis, Colorado State University, Fort Collins, CO, USA.
Kitanidis, P. K.: 1983, 'Statistical estimation of polynomial generalized covariance functions and hydrologic applications', *Water Resources Research* **19**, 909–921.
Legleiter, C. J., Lawrence, R. L., Fonstad, M. A., Marcus, W. A. and Aspinall, R.: 2003, 'Fluvial response a decade after wildfire in the northern Yellowstone ecosystem: A spatially explicit analysis', *Geomorphology* **54**, 119–136.
Levin, S. A.: 1992, 'The problem of pattern and scale in ecology', *Ecology* **73**, 1943–1967.
Little, L. S., Edwards, D. and Porter, D. E.: 1997, 'Kriging in estuaries: As the crow flies, or as the fish swims?', *Journal of Experimental Marine Biology and Ecology* **213**, 1–11.
Mercurio, G., Chaillou, J. C. and Roth, N. E.: 1999, *Guide to Using 1995–1997 Maryland Biological Stream Survey Data*, Maryland Department of Natural Resources, Maryland Department of Natural Resources, Annapolis, MD, USA.
Mulholland, P. J.: 2003, 'Large-scale patterns in dissolved organic carbon concentration, flux, and sources', in: *Aquatic Ecosystems: Interactivity of Dissolved Organic Matter*, Academic Press, San Diego, CA, USA, pp. 139–159.
Neter, J., Kutner, M. H., Wasserman, W. and Nachtsheim, C. J.: 1996, *Applied Linear Statistical Models*, 4th Edition, McGraw-Hill, Boston, MA, USA.
Olden, J. D., Jackson, D. A. and Peres-Neto, P. R.: 2001, 'Spatial isolation and fish communities in drainage lakes', *Oecologia* **127**, 572–585.

Olea, R. A.: 1991, *Geostatistical Glossary and Multilingual Dictionary*, Oxford University Press, New York, NY, USA.

Omernik, J. M.: 1987, 'Ecoregions of the conterminous United States', *Annals of the Association of American Geographers* **77**, 118–125.

Pardo-Iğuzquiza, E.: 1998, 'Maximum likelihood estimation of spatial covariance parameters', *Mathematical Geology* **40**, 95–108.

Peterson, E. E.: 2005, 'Predicting the likelihood of water quality impaired stream segments using landscape-scale data and a hierarchical methodology', PhD Dissertation, Colorado State University, Fort Collins, CO, USA, 293 p.

Poff, N. L.: 1997, 'Landscape filters and species traits: Towards mechanistic understanding and prediction in stream ecology', *Journal of the North American Benthological Society* **16**, 391–409.

Pringle, C. M.: 1991, 'Geothermally modified waters surface at La Selva Biological Station, Costa Rica: Volcanic processes introduce chemical discontinuities into lowland tropical streams', *Biotropica* **40**, 523–529.

Pringle, C. M.: 2001, 'Hydrologic connectivity and the management of biological reserves: A global perspective', *Ecological Applications* **11**, 981–998.

Qualls, R. G. and Haines, B. L.: 1992, 'Biodegradability of dissolved organic matter in forest throughfall, soil solution, and stream water', *Soil Science Society of American Journal* **56**, 578–586.

Rathburn, S. L.: 1998, 'Spatial modeling in irregularly shaped regions: Kriging estuaries', *Environmetrics* **9**, 109–129.

Rousseeuw, P. J. and Leroy, A. M.: 1987, *Robust Regression and Outlier Detection*, John Wiley and Sons, New York, NY, USA.

Shmida, A.: 1984, 'Whittaker plant diversity sampling method', *Israel Journal of Botany* **33**, 41–46.

Smith, K.: 1981, 'The prediction of river water temperatures', *Hydrological Sciences Bulletin* **26**, 19–32.

Stohlgren, T. J., Falkner, M. B. and Schell, L. D.: 1995, 'A Modified-Whittaker nested vegetation sampling method', *Vegetatio* **117**, 113–121.

Torgersen, C. E., Gresswell, R. E. and Bateman, D. S.: 2004, 'Pattern detection in stream networks: quantifying spatial variability in fish distribution', in: *Proceedings of the Second Annual International Symposium on GIS/Spatial Analyses in Fishery and Aquatic Sciences*, Fishery GIS Research Group, Saitama, Japan.

U.S. Environmental Protection Agency (USEPA): 1993, *Regional Environmental Monitoring and Assessment Program*, EPA 625-R-93-012, USEPA, Office of Research and Development, Washington, D.C.

U.S. Geological Survey (USGS): 1999, *Standards for National Hydrography Dataset – Draft*, U.S. Geological Survey and U.S. Environmental Protection Agency, Reston, VA, USA.

Ver Hoef, J. M., Peterson, E. E. and Theobald, D. M.: 2007, 'Some new spatial statistical models for stream networks', *Environmental and Ecological Statistics* **14**, to appear.

Veregin H.: 2000, 'Quantifying positional error induced by line simplification', *International Journal of Geographical Information Science* **14**, 113–130.

Ward, J. V.: 1989, 'The four-dimensional nature of lotic ecosystems', *Journal of the North American Benthological Society* **8**, 2–8.

Wetzel, R. G.: 1992, 'Gradient-dominated ecosystems: Sources and regulatory functions of dissolved organic matter in freshwater ecosystems', *Hydrobiologia* **229**, 181–198.

Yuan, L. L.: 2004, 'Using spatial interpolation to estimate stressor levels in unsampled streams', *Environmental Monitoring and Assessment* **94**, 23–38.